```
************************************************************
*                                                          *
* The CA roles and document type information have been removed from *
* the IDE default display format and the ED field has been added,   *
* effective March 20, 2005.  A new display format, IDERL, is now    *
* available and contains the CA role and document type information. *
*                                                          *
************************************************************
```

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. For more
information enter HELP PROP at an arrow prompt in the file or refer
to the file summary sheet on the web at:
http://www.cas.org/ONLINE/DBSS/registryss.html

FILE CAPLUS

FILE COVERS 1907 - 19 May 2005 VOL 142 ISS 21
FILE LAST UPDATED: 18 May 2005 (20050518/ED)

New CAS Information Use Policies, enter HELP USAGETERMS for details.

   This file contains CAS Registry Numbers for easy and accurate
   substance identification.

FILE MEDLINE
 FILE LAST UPDATED: 18 MAY 2005 (20050518/UP). FILE COVERS 1950 TO DA

 On December 19, 2004, the 2005 MeSH terms were loaded.

 The MEDLINE reload for 2005 is now available. For details enter HELP
 RLOAD at an arrow promt (=>). See also:

    http://www.nlm.nih.gov/mesh/
    http://www.nlm.nih.gov/pubs/techbull/nd04/nd04_mesh.html

 OLDMEDLINE now back to 1950.

 MEDLINE thesauri in the /CN, /CT, and /MN fields incorporate the
 MeSH 2005 vocabulary.

 This file contains CAS Registry Numbers for easy and accurate
 substance identification.

FILE BIOSIS
FILE COVERS 1969 TO DATE.
CAS REGISTRY NUMBERS AND CHEMICAL NAMES (CNs) PRESENT
FROM JANUARY 1969 TO DATE.

                  Searcher   :   Shears   571-272-2528

RECORDS LAST ADDED: 18 May 2005 (20050518/ED)

FILE RELOADED: 19 October 2003.

FILE EMBASE
 FILE COVERS 1974 TO 12 May 2005 (20050512/ED)

 EMBASE has been reloaded. Enter HELP RLOAD for details.

 This file contains CAS Registry Numbers for easy and accurate
 substance identification.

FILE HOME

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2005, 15:47:09 ; Search time 176 Seconds
                                          (without alignments)
                                          113.472 Million cell updates/sec

Title: US-10-629-649A-9
Perfect score: 209
Sequence: 1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39

Scoring table: BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   UniProt_03:*
                  1: uniprot_sprot:*
                  2: uniprot_trembl:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result   Query
   No.   Score Match Length DB ID   Description
----
     1   209   100.0   87 1 EXE4_HELSU   P26349 heloderma s
     2   200   95.7   39 1 EXE3_HELHO   P20394 heloderma h
     3   200   95.7   87 2 Q7SZU6   Q7szu6 heloderma h
     4   119   56.9   266 1 GLU1_XENLA   O42143 xenopus lae
     5   114   54.5   266 2 Q6DIZ4   Q6diz4 xenopus tro
     6   111   53.1   219 1 GLU2_XENLA   O42144 xenopus lae
     7   101   48.3   220 2 Q8UWL9   Q8uwl9 hoplobatrac
     8   99   47.4   31 2 Q7LZN2   Q7lzn2 polyodon sp
     9   97   46.4   31 2 Q7LZN4   Q7lzn4 polyodon sp
    10   97   46.4   71 1 GLUC_ICTPU   P04093 ictalurus p
    11   97   46.4   103 1 GLUC_RANCA   P15438 rana catesb
    12   96   45.9   173 2 Q6RYB9   Q6ryb9 ictalurus p
    13   95   45.5   149 2 Q6RYB2   Q6ryb2 bufo marinu
    14   94   45.0   122 2 Q6RYB8   Q6ryb8 ictalurus p
    15   93   44.5   71 1 GLUC_PIAME   P81880 piaractus m

```
16       93     44.5    121    1   GLUC_CARAU          P79695  carassius a
17       92     44.0     30    1   GLUM_ANGAN          P63294  anguilla an
18       92     44.0     30    1   GLUM_ANGRO          P63295  anguilla ro
19       91     43.5     30    2   Q7LZN3              Q7lzn3  polyodon sp
20       90.5   43.3    178    1   GLU1_ONCMY          Q91971  oncorhynchu
21       90     43.1    120    2   Q6RYB7              Q6ryb7  ictalurus p
22       89     42.6     78    1   GLUC_LEPSP          P09566  lepisosteus
23       89     42.6    153    2   Q6RYB6              Q6ryb6  protopterus
24       88     42.1     66    2   Q788W6              Q788w6  oncorhynchu
25       88     42.1     72    2   Q91409              Q91409  oncorhynchu
26       88     42.1    124    2   Q6RYB1              Q6ryb1  agkistrodon
27       88     42.1    178    1   GLU2_ONCMY          Q91189  oncorhynchu
28       88     42.1    206    1   GLUC_CHICK          P68259  g glucagon
29       87     41.6     45    2   Q6PPF4              Q6ppf4  capra hircu
30       87     41.6    123    2   Q6RYA9              Q6rya9  sebastes ca
31       87     41.6    176    1   GLUC_SHEEP          Q8mj25  o glucagon
32       87     41.6    180    1   GLUC_BOVIN          P01272  b glucagon
33       87     41.6    180    1   GLUC_CANFA          P29794  c glucagon
34       87     41.6    180    1   GLUC_CAVPO          P05110  c glucagon
35       87     41.6    180    1   GLUC_HUMAN          P01275  h glucagon
36       87     41.6    180    1   GLUC_MESAU          P01273  m glucagon
37       87     41.6    180    1   GLUC_MOUSE          P55095  m glucagon .
38       87     41.6    180    1   GLUC_OCTDE          P22890  o glucagon
39       87     41.6    180    1   GLUC_PIG            P01274  s glucagon
40       87     41.6    180    1   GLUC_RAT            P06883  r glucagon
41       86.5   41.4     62    1   GLUC_SCYCA          P09687  scyliorhinu
42       86.5   41.4    160    1   GLU1_PETMA          Q9pur1  petromyzon
43       86     41.1    122    1   GLU2_LOPAM          P04092  lophius ame
44       84     40.2     29    1   GLUC_PLAFE          P23062  platichthys
45       84     40.2     29    1   GLUC_TORMA          P09567  torpedo mar
```

ALIGNMENTS

RESULT 1
EXE4_HELSU
ID   EXE4_HELSU   STANDARD;   PRT;   87 AA.
AC   P26349;
DT   01-MAY-1992 (Rel. 22, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Exendin-4 precursor.
OS   Heloderma suspectum (Gila monster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Helodermatidae;
OC   Heloderma.
OX   NCBI_TaxID=8554;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97172477; PubMed=9020121; DOI=10.1074/jbc.272.7.4335;
RA   Chen Y.E., Drucker D.J.;
RT   "Tissue-specific expression of unique mRNAs that encode proglucagon-
RT   derived peptides or exendin 4 in the lizard.";
RL   J. Biol. Chem. 272:4108-4115(1997).
RN   [2]
RP   SEQUENCE OF 48-86.

```
RC    TISSUE=Venom;
RX    MEDLINE=92218391; PubMed=1313797;
RA    Eng J., Kleinman W.A., Singh L., Singh G., Raufman J.-P.;
RT    "Isolation and characterization of exendin-4, an exendin-3 analogue,
RT    from Heloderma suspectum venom. Further evidence for an exendin
RT    receptor on dispersed acini from guinea pig pancreas.";
RL    J. Biol. Chem. 267:7402-7405(1992).
RN    [3]
RP    STRUCTURE BY NMR OF 48-86.
RX    PubMed=11683627; DOI=10.1021/bi010902s;
RA    Neidigh J.W., Fesinmeyer R.M., Prickett K.S., Andersen N.H.;
RT    "Exendin-4 and glucagon-like-peptide-1: NMR structural comparisons in
RT    the solution and micelle-associated states.";
RL    Biochemistry 40:13188-13200(2001).
CC    -!- FUNCTION: Has a VIP/secretin-like biological activity. Interacts
CC        with the exendin receptor.
CC    -!- SUBCELLULAR LOCATION: Secreted.
CC    -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC    -!- SIMILARITY: Belongs to the glucagon family.
CC    ------------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC    the European Bioinformatics Institute.  There are no  restrictions on  its
CC    use  by  non-profit  institutions as long  as its content  is  in  no  way
CC    modified and this statement is not removed.  Usage  by  and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ------------------------------------------------------------------------
DR    EMBL; U77613; AAB51130.1; -.
DR    PIR; A42486; HWGH4G.
DR    PDB; 1JRJ; NMR; A=48-86.
DR    InterPro; IPR000532; Glucagon.
DR    Pfam; PF00123; Hormone_2; 1.
DR    SMART; SM00070; GLUCA; 1.
DR    PROSITE; PS00260; GLUCAGON; 1.
KW    3D-structure; Amidation; Direct protein sequencing; Glucagon family;
KW    Signal; Toxin.
FT    SIGNAL        1      23       Potential.
FT    PROPEP       24      47
FT    PEPTIDE      48      86       Exendin-4.
FT    MOD_RES      86      86       Serine amide (G-87 provides amide group).
FT    TURN         52      53
FT    HELIX        54      74
FT    TURN         75      76
FT    HELIX        77      79
SQ    SEQUENCE   87 AA;  9479 MW;  656BA6E3D87454A2 CRC64;

      Query Match              100.0%;  Score 209;  DB 1;  Length 87;
      Best Local Similarity    100.0%;  Pred. No. 1.6e-18;
      Matches   39;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39
              |||||||||||||||||||||||||||||||||||||||
Db         48 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 86


RESULT 2
```

EXE3_HELHO
```
ID   EXE3_HELHO      STANDARD;      PRT;     39 AA.
AC   P20394;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Exendin-3.
OS   Heloderma horridum horridum (Mexican beaded lizard).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Helodermatidae;
OC   Heloderma.
OX   NCBI_TaxID=8552;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=91056067; PubMed=1700785;
RA   Eng J., Andrew P.C., Kleinman W.A., Singh L., Raufman J.-P.;
RT   "Purification and structure of exendin-3, a new pancreatic
RT   secretagogue isolated from Heloderma horridum venom.";
RL   J. Biol. Chem. 265:20259-20262(1990).
CC   -!- FUNCTION: Has a VIP/secretin-like biological activity. Interacts
CC       with the exendin receptor.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: Belongs to the glucagon family.
DR   PIR; A23674; HWGH3Z.
DR   HSSP; P26349; 1JRJ.
DR   InterPro; IPR000532; Glucagon.
DR   Pfam; PF00123; Hormone_2; 1.
DR   SMART; SM00070; GLUCA; 1.
DR   PROSITE; PS00260; GLUCAGON; 1.
KW   Amidation; Direct protein sequencing; Glucagon family; Toxin.
FT   MOD_RES      39      39          Serine amide.
SQ   SEQUENCE   39 AA;   4204 MW;   A44251D3A4B1D1B9 CRC64;

   Query Match              95.7%;   Score 200;   DB 1;   Length 39;
   Best Local Similarity    94.9%;   Pred. No. 9e-18;
   Matches    37; Conservative    1; Mismatches    1; Indels    0; Gaps    0;

Qy         1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39
             | :|||||||||||||||||||||||||||||||||||||
Db         1 HSDGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39
```

Search completed: May 18, 2005, 15:57:22
Job time : 177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2005, 15:47:50 ; Search time 39 Seconds
                                         (without alignments)
                                         96.217 Million cell updates/sec

Title: US-10-629-649A-9
Perfect score: 209
Sequence: 1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_79:*
              1: pir1:*
              2: pir2:*
              3: pir3:*
              4: pir4:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                             SUMMARIES
                    %
Result   Query
   No.   Score   Match Length DB   ID   Description
----
     1   209   100.0   39 1   HWGH4G   exendin-4 - Gila m
     2   200   95.7   39 1   HWGH3Z   exendin-3 - Mexica
     3   99   47.4   31 2   S44472   glucagon G2 - Nort
     4   97   46.4   31 2   S44471   glucagon G1 - Nort
     5   97   46.4   101 1   GCFGB   glucagon precursor
     6   94   45.0   63 1   GCIDC   glucagon precursor
     7   92   44.0   30 2   C61125   glucagon-like pept
     8   92   44.0   30 2   B61125   glucagon-like pept
     9   91   43.5   30 2   S44473   glucagon-like pept
    10   90.5   43.3   178 2   I51058   glucagon I precurs
    11   89   42.6   72 1   GCGXA   glucagon precursor
    12   88   42.1   66 2   I51093   glucagon - chinook
    13   88   42.1   151 1   GCCH   glucagon precursor

```
14      88     42.1    178    2   I51057                     glucagon II precur
15      88     42.1    206    2   I51301                     proglucagon - chic
16      87     41.6     29    1   GCDF                       glucagon - smaller
17      87     41.6    158    1   GCPG                       glucagon precursor
18      87     41.6    180    1   GCBO                       glucagon precursor
19      87     41.6    180    1   GCHY                       glucagon precursor
20      87     41.6    180    1   GCGP                       glucagon precursor
21      87     41.6    180    1   GCHU                       glucagon precursor
22      87     41.6    180    1   GCRT                       glucagon precursor
23      87     41.6    180    1   GCRTDU                     glucagon precursor
24      87     41.6    180    2   A57294                     glucagon precursor
25      86     41.1    122    1   GCAF2                      glucagon 2 precurs
26      84     40.2     29    1   GCFLE                      glucagon - Europea
27      84     40.2     29    2   S07211                     glucagon - marbled
28      84     40.2     29    2   A61135                     glucagon - bigeye
29      84     40.2     87    1   GCFIS                      glucagon precursor
30      83     39.7     60    1   GCONC                      glucagon precursor
31      81     38.8     29    1   GCCB                       glucagon - Chinchi
32      81     38.8    124    1   GCAF                       glucagon 1 precurs
33      80     38.3     29    2   C39258                     glucagon - common
34      80     38.3     29    2   A91742                     glucagon - Arabian
35      80     38.3     29    2   A91741                     glucagon - rabbit
36      80     38.3     69    1   GCDG69                     glucagon-69 - dog
37      79     37.8     29    2   C60840                     glucagon I - Europ
38      75     35.9     29    1   GCEN                       glucagon - elephan
39      75     35.9     29    1   GCOPV                      glucagon - North A
40      75     35.9     29    2   A91740                     glucagon - turkey
41      74     35.4     29    1   A61583                     glucagon - ostrich
42      74     35.4     29    1   GCDK                       glucagon - duck
43      74     35.4     29    1   GCTTS                      glucagon - slider
44      74     35.4     29    2   S39018                     glucagon - bowfin
45    71.5     34.2     36    2   D60840                     glucagon II - Euro
```

ALIGNMENTS

RESULT 1
HWGH4G
exendin-4 - Gila monster
C;Species: Heloderma suspectum (Gila monster)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: A42486
R;Eng, J.; Kleinman, W.A.; Singh, L.; Singh, G.; Raufman, J.P.
J. Biol. Chem. 267, 7402-7405, 1992
A;Title: Isolation and characterization of exendin-4, an exendin-3 analogue,
from Heloderma suspectum venom. Further evidence for an exendin receptor on
dispersed acini from guinea pig pancreas.
A;Reference number: A42486; MUID:92218391; PMID:1313797
A;Accession: A42486
A;Molecule type: protein
A;Residues: 1-39 <ENG>
A;Cross-references: UNIPROT:P26349
C;Comment: Exendin-4 does not stimulate amylase secretion by pancreatic acinar
cells.
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; duplication; venom

F;39/Modified site: amidated carboxyl end (Ser) #status experimental

      Query Match   100.0%; Score 209; DB 1; Length 39;
      Best Local Similarity   100.0%; Pred. No. 3.4e-19;
      Matches   39; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39
            ||||||||||||||||||||||||||||||||||||||||
Db   1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39

RESULT 2
HWGH3Z
exendin-3 - Mexican beaded lizard
C;Species: Heloderma horridum (Mexican beaded lizard)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
C;Accession: A23674
R;Eng, J.; Andrews, P.C.; Kleinman, W.A.; Singh, L.; Raufman, J.P.
J. Biol. Chem. 265, 20259-20262, 1990
A;Title: Purification and structure of exendin-3, a new pancreatic secretagogue
isolated from Heloderma horridum venom.
A;Reference number: A23674; MUID:91056067; PMID:1700785
A;Accession: A23674
A;Molecule type: protein
A;Residues: 1-39 <ENG>
A;Cross-references: UNIPROT:P20394
C;Comment: Exendins are venom components that are thought to bind to receptors
for vasoactive intestinal peptide and/or secretin on pancreatic acinar cells and
to activate adenylate cyclase, resulting in secretion of amylase.
C;Superfamily: glucagon
C;Keywords: amidated carboxyl end; duplication; secretagogue; venom
F;39/Modified site: amidated carboxyl end (Ser) #status experimental

      Query Match   95.7%; Score 200; DB 1; Length 39;
      Best Local Similarity   94.9%; Pred. No. 4.4e-18;
      Matches   37; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

Qy   1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39
            | :|||||||||||||||||||||||||||||||||||||
Db   1 HSDGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39

Search completed: May 18, 2005, 15:58:09
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2005, 15:57:31 ; Search time 138 Seconds
                                         (without alignments)
                                         94.535 Million cell updates/sec

Title: US-10-629-649A-9
Perfect score: 209
Sequence: 1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published_Applications_AA:*
                 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
                 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
                 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
                 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
                 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
                 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
                 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
                 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
                 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                            SUMMARIES
          %

Result No.	Score	Query Match	Length	DB	ID	Description
1	209	100.0	39	9	US-09-876-388-12	Sequence 12, Appl
2	209	100.0	39	9	US-09-851-738-9	Sequence 9, Appli
3	209	100.0	39	9	US-09-805-507-9	Sequence 9, Appli
4	209	100.0	39	9	US-09-859-804-9	Sequence 9, Appli
5	209	100.0	39	9	US-09-003-869-2	Sequence 2, Appli
6	209	100.0	39	9	US-09-982-978-9	Sequence 9, Appli
7	209	100.0	39	9	US-09-953-021B-9	Sequence 9, Appli
8	209	100.0	39	10	US-09-756-690A-2	Sequence 2, Appli
9	209	100.0	39	14	US-10-091-258-9	Sequence 9, Appli
10	209	100.0	39	14	US-10-157-224A-2	Sequence 2, Appli
11	209	100.0	39	14	US-10-187-051-2	Sequence 2, Appli
12	209	100.0	39	14	US-10-055-259-9	Sequence 9, Appli
13	209	100.0	39	14	US-10-287-892-12	Sequence 12, Appl
14	209	100.0	39	14	US-10-288-340-12	Sequence 12, Appl
15	209	100.0	39	14	US-10-097-230-4	Sequence 4, Appli
16	209	100.0	39	15	US-10-322-839-9	Sequence 9, Appli
17	209	100.0	39	15	US-10-342-014-2	Sequence 2, Appli
18	209	100.0	39	15	US-10-433-108-10	Sequence 10, Appl
19	209	100.0	39	15	US-10-408-262B-1	Sequence 1, Appli
20	209	100.0	39	16	US-10-291-226-102	Sequence 102, App
21	209	100.0	39	16	US-10-722-733-12	Sequence 12, Appl
22	209	100.0	39	16	US-10-723-099-12	Sequence 12, Appl
23	209	100.0	39	16	US-10-704-409-1	Sequence 1, Appli
24	209	100.0	39	16	US-10-726-734-1	Sequence 1, Appli
25	209	100.0	39	17	US-10-718-071-30	Sequence 30, Appl
26	209	100.0	39	17	US-10-850-055-4	Sequence 4, Appli
27	209	100.0	39	17	US-10-850-055-21	Sequence 21, Appl
28	209	100.0	39	17	US-10-903-767-9	Sequence 9, Appli
29	209	100.0	39	17	US-10-656-093C-2	Sequence 2, Appli
30	209	100.0	39	17	US-10-892-981A-4	Sequence 4, Appli
31	209	100.0	39	17	US-10-895-909-2	Sequence 2, Appli
32	209	100.0	39	17	US-10-964-887-2	Sequence 2, Appli
33	209	100.0	39	17	US-10-966-337-2	Sequence 2, Appli
34	209	100.0	39	18	US-11-007-938-9	Sequence 9, Appli
35	209	100.0	40	9	US-09-876-388-18	Sequence 18, Appl
36	209	100.0	40	9	US-09-876-388-31	Sequence 31, Appl
37	209	100.0	40	9	US-09-876-388-32	Sequence 32, Appl
38	209	100.0	40	14	US-10-287-892-18	Sequence 18, Appl
39	209	100.0	40	14	US-10-287-892-31	Sequence 31, Appl
40	209	100.0	40	14	US-10-287-892-32	Sequence 32, Appl
41	209	100.0	40	14	US-10-288-340-18	Sequence 18, Appl
42	209	100.0	40	14	US-10-288-340-31	Sequence 31, Appl
43	209	100.0	40	14	US-10-288-340-32	Sequence 32, Appl
44	209	100.0	40	16	US-10-291-226-138	Sequence 138, App
45	209	100.0	40	16	US-10-291-226-151	Sequence 151, App

ALIGNMENTS

RESULT 1
US-09-876-388-12
; Sequence 12, Application US/09876388
; Patent No. US20020049153A1

```
;  GENERAL INFORMATION:
;   APPLICANT: Bridon, Dominique P.
;   APPLICANT:  L'Archeveque, Benoit
;   APPLICANT:  Ezrin, Alan M.
;   APPLICANT:  Holmes, Darren L.
;   APPLICANT:  Leblanc, Anouk
;   APPLICANT:  St. Pierre, Serge
;   TITLE OF INVENTION: LONG LASTING INSULINOTROPIC PEPTIDES
;   FILE REFERENCE: 500862001610
;   CURRENT APPLICATION NUMBER: US/09/876,388
;   CURRENT FILING DATE:  2001-09-24
;   PRIOR APPLICATION NUMBER: 09/623,618
;   PRIOR FILING DATE: 2000-09-05
;   PRIOR APPLICATION NUMBER: PCT/US00/13563
;   PRIOR FILING DATE: 2000-05-17
;   PRIOR APPLICATION NUMBER: 60/159,783
;   PRIOR FILING DATE: 1999-10-15
;   PRIOR APPLICATION NUMBER: 60/134,406
;   PRIOR FILING DATE: 1999-05-17
;   NUMBER OF SEQ ID NOS: 35
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 39
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Peptide
US-09-876-388-12

   Query Match              100.0%;  Score 209;  DB 9;  Length 39;
   Best Local Similarity    100.0%;  Pred. No. 2e-19;
   Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39
            |||||||||||||||||||||||||||||||||||||||
Db        1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39


Search completed: May 18, 2005, 16:10:40
Job time : 140 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 18, 2005, 15:54:30 ; Search time 43 Seconds
                                          (without alignments)
                                   67.705 Million cell updates/sec

Title: US-10-629-649A-9
Perfect score: 209
Sequence: 1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
                 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result   Query
  No.   Score Match Length DB ID   Description
----
     1   209 100.0   39 1 US-08-066-480-2   Sequence 2, Appli
     2   209 100.0   39 3 US-09-302-596-9   Sequence 9, Appli
     3   209 100.0   39 3 US-09-623-618B-12   Sequence 12, Appl
     4   209 100.0   39 3 US-09-333-415-9   Sequence 9, Appli
     5   209 100.0   39 4 US-09-303-016-9   Sequence 9, Appli
     6   209 100.0   39 4 US-09-323-867A-2   Sequence 2, Appli
     7   209 100.0   39 4 US-09-657-332A-12   Sequence 12, Appl
     8   209 100.0   39 4 US-09-614-847-102   Sequence 102, App
     9   209 100.0   39 4 US-09-805-507-9   Sequence 9, Appli
    10   209 100.0   39 4 US-09-876-388-12   Sequence 12, Appl
    11   209 100.0   39 4 US-09-622-105-2   Sequence 2, Appli

```
12     209    100.0    39    4    US-09-859-804-9           Sequence 9, Appli
13     209    100.0    39    4    US-10-287-892-12          Sequence 12, Appl
14     209    100.0    40    3    US-09-623-618B-18         Sequence 18, Appl
15     209    100.0    40    3    US-09-623-618B-31         Sequence 31, Appl
16     209    100.0    40    3    US-09-623-618B-32         Sequence 32, Appl
17     209    100.0    40    4    US-09-657-332A-18         Sequence 18, Appl
18     209    100.0    40    4    US-09-657-332A-31         Sequence 31, Appl
19     209    100.0    40    4    US-09-657-332A-32         Sequence 32, Appl
20     209    100.0    40    4    US-09-614-847-138         Sequence 138, App
21     209    100.0    40    4    US-09-614-847-151         Sequence 151, App
22     209    100.0    40    4    US-09-876-388-18          Sequence 18, Appl
23     209    100.0    40    4    US-09-876-388-31          Sequence 31, Appl
24     209    100.0    40    4    US-09-876-388-32          Sequence 32, Appl
25     209    100.0    40    4    US-10-287-892-18          Sequence 18, Appl
26     209    100.0    40    4    US-10-287-892-31          Sequence 31, Appl
27     209    100.0    40    4    US-10-287-892-32          Sequence 32, Appl
28     209    100.0    45    4    US-09-614-847-92          Sequence 92, Appl
29     209    100.0    47    4    US-09-614-847-100         Sequence 100, App
30     209    100.0    87    4    US-09-019-172-2           Sequence 2, Appli
31     208     99.5    39    4    US-09-323-867A-25         Sequence 25, Appl
32     206     98.6    39    4    US-09-323-867A-10         Sequence 10, Appl
33     206     98.6    39    4    US-09-323-867A-14         Sequence 14, Appl
34     206     98.6    39    4    US-09-323-867A-18         Sequence 18, Appl
35     206     98.6    39    4    US-09-323-867A-29         Sequence 29, Appl
36     205     98.1    38    4    US-09-323-867A-62         Sequence 62, Appl
37     205     98.1    38    4    US-09-614-847-104         Sequence 104, App
38     205     98.1    38    4    US-09-614-847-129         Sequence 129, App
39     205     98.1    38    4    US-09-622-105-28          Sequence 28, Appl
40     205     98.1    39    4    US-09-323-867A-13         Sequence 13, Appl
41     205     98.1    39    4    US-09-323-867A-16         Sequence 16, Appl
42     205     98.1    39    4    US-09-323-867A-19         Sequence 19, Appl
43     205     98.1    39    4    US-09-614-847-107         Sequence 107, App
44     205     98.1    39    4    US-09-614-847-126         Sequence 126, App
45     205     98.1    44    4    US-09-614-847-91          Sequence 91, Appl
```

ALIGNMENTS

RESULT 1
US-08-066-480-2
; Sequence 2, Application US/08066480
; Patent No. 5424286
; GENERAL INFORMATION:
;   APPLICANT: Eng, John
;   TITLE OF INVENTION: Pharmaceutical Compositions And Use of
;   TITLE OF INVENTION: Exendin-3 and Exendin-4 for Treatment of Diabetes
Mellitus
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Allegretti & Witcoff, Ltd.
;   STREET: 10 S. Wacker Drive
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60606
;   COMPUTER READABLE FORM:

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/066,480
;       FILING DATE:  24-MAR-1993
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McDonnell, John J
;       REGISTRATION NUMBER:  26,949
;       REFERENCE/DOCKET NUMBER:  93,084
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-715-1000
;       TELEFAX:  312-715-1234
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  39 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  Peptide
;       LOCATION:  1..39
;       OTHER INFORMATION:  /label= Exendin-4
US-08-066-480-2

   Query Match            100.0%;  Score 209;  DB 1;  Length 39;
   Best Local Similarity   100.0%;  Pred. No. 5.5e-20;
   Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39
            |||||||||||||||||||||||||||||||||||||||
Db        1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39


Search completed: May 18, 2005, 16:08:14
Job time : 44 secs
```

OM protein - protein search, using sw model

Run on: May 18, 2005, 15:46:14 ; Search time 74 Seconds
                                          (without alignments)
                                          203.833 Million cell updates/sec

Title: US-10-629-649A-9
Perfect score: 209
Sequence: 1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   A_Geneseq_16Dec04:*
                 1: geneseqp1980s:*
                 2: geneseqp1990s:*
                 3: geneseqp2000s:*
                 4: geneseqp2001s:*
                 5: geneseqp2002s:*
                 6: geneseqp2003as:*
                 7: geneseqp2003bs:*
                 8: geneseqp2004s:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                             SUMMARIES

		% Query				
Result No.	Score	Match	Length	DB	ID	Description
1	209	100.0	39	2	AAR80546	Aar80546 Heloderma
2	209	100.0	39	2	AAW47609	Aaw47609 Gila mons
3	209	100.0	39	2	AAW61770	Aaw61770 Exendin-4
4	209	100.0	39	2	AAY03718	Aay03718 Amino aci
5	209	100.0	39	2	AAY31502	Aay31502 Exendin-4
6	209	100.0	39	2	ADI24854	Adi24854 Exendin-4
7	209	100.0	39	3	AAY78957	Aay78957 Extendin-
8	209	100.0	39	3	AAY94011	Aay94011 Amino aci
9	209	100.0	39	3	AAB11306	Aab11306 exendin a

```
10      209   100.0   39  3   AAB11307            Aab11307 exendin a
11      209   100.0   39  3   AAB11305            Aab11305 exendin a
12      209   100.0   39  3   AAB11282            Aab11282 H. suspec
13      209   100.0   39  3   AAB11308            Aab11308 exendin a
14      209   100.0   39  3   AAB52872            Aab52872 Gila mons
15      209   100.0   39  3   AAB52857            Aab52857 Extendin-
16      209   100.0   39  3   AAB52840            Aab52840 Extendin-
17      209   100.0   39  3   AAB52858            Aab52858 Extendin-
18      209   100.0   39  3   AAB52841            Aab52841 Extendin-
19      209   100.0   39  4   AAB69971            Aab69971 Exendin-4
20      209   100.0   39  4   AAB36421            Aab36421 Gila mons
21      209   100.0   39  4   AAB48801            Aab48801 Exendin-4
22      209   100.0   39  4   AAE08346            Aae08346 Heloderma
23      209   100.0   39  4   AAB64182            Aab64182 Gila mons
24      209   100.0   39  4   AAU07380            Aau07380 Glucagon-
25      209   100.0   39  4   AAB91178            Aab91178 Pancreati
26      209   100.0   39  4   AAB91175            Aab91175 Pancreati
27      209   100.0   39  4   AAB91191            Aab91191 Pancreati
28      209   100.0   39  4   AAG70462            Aag70462 Exendin-4
29      209   100.0   39  4   AAB85927            Aab85927 Gila mons
30      209   100.0   39  4   AAB60254            Aab60254 Gila mons
31      209   100.0   39  4   AAB36434            Aab36434 Gila mons
32      209   100.0   39  5   AAE14427            Aae14427 Gila mons
33      209   100.0   39  5   ABB83059            Abb83059 Glucagon
34      209   100.0   39  5   ABG71259            Abg71259 Gila mons
35      209   100.0   39  5   ABB07151            Abb07151 Gila mons
36      209   100.0   39  6   ABU66208            Abu66208 Gila mons
37      209   100.0   39  6   ABU91974            Abu91974 Glia mons
38      209   100.0   39  6   AAE30913            Aae30913 Exendin-4
39      209   100.0   39  6   AAE30938            Aae30938 Human GLP
40      209   100.0   39  6   ABP58578            Abp58578 Mexican b
41      209   100.0   39  6   AAO19593            Aao19593 Gila mons
42      209   100.0   39  6   ADA44872            Ada44872 Gila mons
43      209   100.0   39  7   ADB84205            Adb84205 Gila mons
44      209   100.0   39  7   ADD02757            Add02757 Exendin-4
45      209   100.0   39  7   ADG73289            Adg73289 Glucagon-
```

                              ALIGNMENTS

RESULT 1
AAR80546
ID   AAR80546 standard; peptide; 39 AA.
XX
AC   AAR80546;
XX
DT   27-FEB-1996 (first entry)
XX
DE   Heloderma suspectum exendin-4.
XX
KW   Exendin-4; diabetes mellitus; hyperglycaemia; insulinotropic peptide.
XX
OS   Heloderma suspectum.
XX
PN   US5424286-A.
XX

```
PD    13-JUN-1995.
XX
PF    24-MAY-1993;    93US-00066480.
XX
PR    24-MAY-1993;    93US-00066480.
XX
PA    (ENGJ/) ENG J.
XX
PI    Eng J;
XX
DR    WPI; 1995-262627/34.
XX
PT    Stimulating/inhibiting insulin release with exendin polypeptide(s) - for
PT    treating diabetes mellitus and preventing hyperglycaemia.
XX
PS    ·Claim 6; Col 13-14; 17pp; English.
XX
CC    AAR80546 is Heloderma suspectum exendin-4. It is an insulinotropic
CC    peptide, and can therefore be used in the treatment of diabetes mellitus
CC    (types I or II), and for the prevention of hyperglycaemia. It normalises
CC    hyperglycaemia through glucose-dependent and insulin-(in)dependent
CC    mechanisms
XX
SQ    Sequence 39 AA;

   Query Match              100.0%;   Score 209;   DB 2;   Length 39;
   Best Local Similarity    100.0%;   Pred. No. 6.2e-19;
   Matches    39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39
             |||||||||||||||||||||||||||||||||||||||
Db         1 HGEGTFTSDLSKQMEEEAVRLFIEWLKNGGPSSGAPPPS 39


Search completed: May 18, 2005, 15:54:22
Job time : 78 secs
```